

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- 5 (i) APPLICANT: Blank, Gregory S.  
Narindray, Daljit S.  
Zapata, Gerardo A.
- (ii) TITLE OF INVENTION: Protein Recovery
- 10 (iii) NUMBER OF SEQUENCES: 7
- (iv) CORRESPONDENCE ADDRESS:  
15 (A) ADDRESSEE: Genentech, Inc.  
(B) STREET: 1 DNA Way  
(C) CITY: South San Francisco  
(D) STATE: California  
(E) COUNTRY: USA  
20 (F) ZIP: 94080
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
25 (D) SOFTWARE: WinPatin (Genentech)
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: to be assigned  
(B) FILING DATE: 21-Jan-2004  
30 (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: 09/940166  
(B) FILING DATE: 27-Aug-2001  
35
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: 09/460587  
(B) FILING DATE: 14 Dec-1999
- 40 (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: 09/097309  
(B) FILING DATE: 12-Jun-1998
- (vii) PRIOR APPLICATION DATA:  
45 (A) APPLICATION NUMBER: 60/050951  
(B) FILING DATE: 13-JUN-1997
- (viii) ATTORNEY/AGENT INFORMATION:  
50 (A) NAME: Lee K. Tan  
(B) REGISTRATION NUMBER: 39447  
(C) REFERENCE/DOCKET NUMBER: P1105R1D1C2
- (ix) TELECOMMUNICATION INFORMATION:  
55 (A) TELEPHONE: 650/225-4462  
(B) TELEFAX: 650/952-9881
- (2) INFORMATION FOR SEQ ID NO:1:
- 60 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 241 amino acids  
(B) TYPE: Amino Acid  
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

	Glu	Val	Gln	Leu	Val	Glu	Ser	Gly	Gly	Gly	Leu	Val	Gln	Pro	Gly	
	1				5					10					15	
5	Gly	Ser	Leu	Arg	Leu	Ser	Cys	Ala	Thr	Ser	Gly	Tyr	Thr	Phe	Thr	
					20					25					30	
	Glu	Tyr	Thr	Met	His	Trp	Met	Arg	Gln	Ala	Pro	Gly	Lys	Gly	Leu	
10					35					40					45	
	Glu	Trp	Val	Ala	Gly	Ile	Asn	Pro	Lys	Asn	Gly	Gly	Thr	Ser	His	
					50					55					60	
15	Asn	Gln	Arg	Phe	Met	Asp	Arg	Phe	Thr	Ile	Ser	Val	Asp	Lys	Ser	
					65					70					75	
	Thr	Ser	Thr	Ala	Tyr	Met	Gln	Met	Asn	Ser	Leu	Arg	Ala	Glu	Asp	
					80					85					90	
20	Thr	Ala	Val	Tyr	Tyr	Cys	Ala	Arg	Trp	Arg	Gly	Leu	Asn	Tyr	Gly	
					95					100					105	
	Phe	Asp	Val	Arg	Tyr	Phe	Asp	Val	Trp	Gly	Gln	Gly	Thr	Leu	Val	
25					110					115					120	
	Thr	Val	Ser	Ser	Ala	Ser	Thr	Lys	Gly	Pro	Ser	Val	Phe	Pro	Leu	
					125					130					135	
30	Ala	Pro	Ser	Ser	Lys	Ser	Thr	Ser	Gly	Gly	Thr	Ala	Ala	Leu	Gly	
					140					145					150	
	Cys	Leu	Val	Lys	Asp	Tyr	Phe	Pro	Glu	Pro	Val	Thr	Val	Ser	Trp	
					155					160					165	
35	Asn	Ser	Gly	Ala	Leu	Thr	Ser	Gly	Val	His	Thr	Phe	Pro	Ala	Val	
					170					175					180	
	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser	Leu	Ser	Ser	Val	Val	Thr	Val	
40					185					190					195	
	Pro	Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr	Tyr	Ile	Cys	Asn	Val	Asn	
					200					205					210	
45	His	Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys	Lys	Val	Glu	Pro	Lys	
					215					220					225	
	Ser	Cys	Asp	Lys	Thr	His	Thr	Cys	Pro	Pro	Cys	Pro	Ala	Pro	Glu	
					230					235					240	
50	Leu															
	241															

(2) INFORMATION FOR SEQ ID NO:2:

55

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 214 amino acids
  - (B) TYPE: Amino Acid
  - (D) TOPOLOGY: Linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

	Asp	Ile	Gln	Met	Thr	Gln	Ser	Pro	Ser	Ser	Leu	Ser	Ala	Ser	Val	
	1				5					10					15	
5	Gly	Asp	Arg	Val	Thr	Ile	Thr	Cys	Arg	Ala	Ser	Gln	Asp	Ile	Asn	
					20					25					30	
	Asn	Tyr	Leu	Asn	Trp	Tyr	Gln	Gln	Lys	Pro	Gly	Lys	Ala	Pro	Lys	
					35					40					45	
10	Leu	Leu	Ile	Tyr	Tyr	Thr	Ser	Thr	Leu	His	Ser	Gly	Val	Pro	Ser	
					50					55					60	
	Arg	Phe	Ser	Gly	Ser	Gly	Ser	Gly	Thr	Asp	Tyr	Thr	Leu	Thr	Ile	
					65					70					75	
15	Ser	Ser	Leu	Gln	Pro	Glu	Asp	Phe	Ala	Thr	Tyr	Tyr	Cys	Gln	Gln	
					80					85					90	
	Gly	Asn	Thr	Leu	Pro	Pro	Thr	Phe	Gly	Gln	Gly	Thr	Lys	Val	Glu	
20					95					100					105	
	Ile	Lys	Arg	Thr	Val	Ala	Ala	Pro	Ser	Val	Phe	Ile	Phe	Pro	Pro	
					110					115					120	
25	Ser	Asp	Glu	Gln	Leu	Lys	Ser	Gly	Thr	Ala	Ser	Val	Val	Cys	Leu	
					125					130					135	
	Leu	Asn	Asn	Phe	Tyr	Pro	Arg	Glu	Ala	Lys	Val	Gln	Trp	Lys	Val	
					140					145					150	
30	Asp	Asn	Ala	Leu	Gln	Ser	Gly	Asn	Ser	Gln	Glu	Ser	Val	Thr	Glu	
					155					160					165	
	Gln	Asp	Ser	Lys	Asp	Ser	Thr	Tyr	Ser	Leu	Ser	Ser	Thr	Leu	Thr	
35					170					175					180	
	Leu	Ser	Lys	Ala	Asp	Tyr	Glu	Lys	His	Lys	Val	Tyr	Ala	Cys	Glu	
					185					190					195	
40	Val	Thr	His	Gln	Gly	Leu	Ser	Ser	Pro	Val	Thr	Lys	Ser	Phe	Asn	
					200					205					210	
	Arg	Gly	Glu	Cys												
					214											

45

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 36 amino acids  
 (B) TYPE: Amino Acid  
 (D) TOPOLOGY: Linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

55	Leu	Gly	Gly	Arg	Met	Lys	Gln	Leu	Glu	Asp	Lys	Val	Glu	Glu	Leu	
	1				5					10					15	
	Leu	Ser	Lys	Asn	Tyr	His	Leu	Glu	Asn	Glu	Val	Ala	Arg	Leu	Lys	
					20					25					30	
60	Lys	Leu	Val	Gly	Glu	Arg										
					35	36										

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 7 amino acids  
 (B) TYPE: Amino Acid  
 (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Leu Xaa Xaa Xaa Xaa Xaa Xaa  
 1 5 7

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2143 base pairs  
 (B) TYPE: Nucleic Acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GAATTCAACT TCTCCATACT TTGGATAAGG AAATACAGAC ATGAAAAATC 50  
 TCATTGCTGA GTTGTATTAT AAGCTTTGGA GATTATCGTC ACTGCAATGC 100  
 TTCGCAATAT GCGCAAAAT GACCAACAGC GGTGATTGA TCAGGTAGAG 150  
 GGGGCGCTGT ACGAGGTAAA GCCCGATGCC AGCATTCCTG ACGACGATAC 200  
 GGAGCTGCTG CCGGATTACG TAAAGAAGTT ATTGAAGCAT CCTCGTCAGT 250  
 AAAAAGTTAA TCTTTTCAAC AGCTGTCATA AAGTTGTCAC GGCCGAGACT 300  
 TATAGTCGCT TTGTTTTTAT TTTTAAATGT ATTTGTAATC AGAATTTCGAG 350  
 CTCGCCGGGG ATCCTCTAGA GGTGAGGTG ATTTTATGAA AAAGAATATC 400  
 GCATTTCTTC TTGCATCTAT GTTCGTTTTT TCTATTGCTA CAAACGCGTA 450  
 CGCTGATATC CAGATGACCC AGTCCCCGAG CTCCCTGTCC GCCTCTGTGG 500  
 GCGATAGGGT CACCATCACC TGTCGTGCCA GTCAGGACAT CAACAATTAT 550  
 CTGAAGTGGT ATCAACAGAA ACCAGGAAAA GCTCCGAAAC TACTGATTTA 600  
 CTATACCTCC ACCCTCCACT CTGGAGTCCC TTCTCGCTTC TCTGGTTCTG 650  
 GTTCTGGGAC GGATTACACT CTGACCATCA GCAGTCTGCA ACCGGAGGAC 700  
 TTCGCAACTT ATTACTGTCA GCAAGGTAAT ACTCTGCCGC CGACGTTCCG 750  
 ACAGGGCAGC AAGGTGGAGA TCAAACGAAC TGTGGCTGCA CCATCTGTCT 800  
 TCATCTTCCC GCCATCTGAT GAGCAGTTGA AATCTGGAAC TGCCTCTGTT 850  
 GTGTGCCTGC TGAATAACTT CTATCCCAGA GAGGCCAAAG TACAGTGGAA 900  
 GGTGGATAAC GCCCTCCAAT CGGGTAACTC CCAGGAGAGT GTCACAGAGC 950  
 AGGACAGCAA GGACAGCACC TACAGCCTCA GCAGCACCCCT GACGCTGAGC 1000

AAAGCAGACT ACGAGAAACA CAAAGTCTAC GCCTGCGAAG TCACCCATCA 1050  
 GGGCCTGAGC TCGCCCGTCA CAAAGAGCTT CAACAGGGGA GAGTGTAAAG 1100  
 5 CTGATCCTCT ACGCCGGACG CATCGTGGCG CTAGTACGCA AGTTCACGTA 1150  
 AAAACGGTAT CTAGAGGTTG AGGTGATTTT ATGAAAAAGA ATATCGCATT 1200  
 10 TCTTCTTGCA TCTATGTTCTG TTTTCTCTAT TGCTACAAAC GCGTACGCTG 1250  
 AGGTTTCAGCT GGTGGAGTCT GGCAGTGGCC TGGTGCAGCC AGGGGGCTCA 1300  
 CTCCGTTTGT CCTGTGCAAC TTCTGGCTAC ACCTTTACCG AATACACTAT 1350  
 15 GCACTGGATG CGTCAGGCCC CGGTAAGGG CCTGGAATGG GTTGCAGGGA 1400  
 TTAATCCTAA AAACGGTGGT ACCAGCCACA ACCAGAGGT CATGGACCGT 1450  
 TTCACTATAA GCGTAGATAA ATCCACCAGT ACAGCCTACA TGCAAATGAA 1500  
 20 CAGCCTGCGT GCTGAGGACA CTGCCGTCTA TTATTGTGCT AGATGGCGAG 1550  
 GCCTGAACTA CGGCTTTGAC GTCCGTTATT TTGACGCTG GGGTCAAGGA 1600  
 25 ACCCTGGTCA CCGTCTCCTC GGCCTCCACC AAGGGCCCAT CCGTCTTCCC 1650  
 CCTGGCACCC TCCTCCAAGA GCACCTCTGG GGGCACAGCG GCCCTGGGCT 1700  
 GCCTGGTCAA GGACTACTTC CCCGAACCGG TGACGGTGTC GTGGAAGTCA 1750  
 30 GCGGCCCTGA CCAGCGGCGT GCACACCTTC CCGGCTGTCC TACAGTCCTC 1800  
 AGGACTCTAC TCCCTCAGCA GCGTGGTGAC CGTGCCCTCC AGCAGCTTGG 1850  
 35 GCACCCAGAC CTACATCTGC AACGTGAATC ACAAGCCCAG CAACACCAAG 1900  
 GTCGACAAGA AAGTTGAGCC CAAATCTTGT GACAAAATC ACACATGCCC 1950  
 GCCGTGCCCA GCACCAGAAC TGCTGGGCGG CCGCATGAAA CAGCTAGAGG 2000  
 40 ACAAGGTCGA AGAGCTACTC TCCAAGAACT ACCACCTAGA GAATGAAGTG 2050  
 GCAAGACTCA AAAAGCTTGT CGGGGAGCGC TAAGCATGCG ACGGCCCTAG 2100  
 45 AGTCCCTAAC GCTCGGTTGC CGCCGGGCGT TTTTATTGT TAA 2143

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:  
 50 (A) LENGTH: 237 amino acids  
 (B) TYPE: Amino Acid  
 (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

55 Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe  
 -23 -20 -15 -10  
 Ser Ile Ala Thr Asn Ala Tyr Ala Asp Ile Gln Met Thr Gln Ser  
 60 -5 1 5  
 Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr  
 10 15 20

	Cys	Arg	Ala	Ser	Gln	Asp	Ile	Asn	Asn	Tyr	Leu	Asn	Trp	Tyr	Gln	
			25					30					35			
5	Gln	Lys	Pro	Gly	Lys	Ala	Pro	Lys	Leu	Leu	Ile	Tyr	Tyr	Thr	Ser	
			40					45					50			
	Thr	Leu	His	Ser	Gly	Val	Pro	Ser	Arg	Phe	Ser	Gly	Ser	Gly	Ser	
			55					60					65			
10	Gly	Thr	Asp	Tyr	Thr	Leu	Thr	Ile	Ser	Ser	Leu	Gln	Pro	Glu	Asp	
			70					75					80			
	Phe	Ala	Thr	Tyr	Tyr	Cys	Gln	Gln	Gly	Asn	Thr	Leu	Pro	Pro	Thr	
15			85					90					95			
	Phe	Gly	Gln	Gly	Thr	Lys	Val	Glu	Ile	Lys	Arg	Thr	Val	Ala	Ala	
			100					105					110			
20	Pro	Ser	Val	Phe	Ile	Phe	Pro	Pro	Ser	Asp	Glu	Gln	Leu	Lys	Ser	
			115					120					125			
	Gly	Thr	Ala	Ser	Val	Val	Cys	Leu	Leu	Asn	Asn	Phe	Tyr	Pro	Arg	
			130					135					140			
25	Glu	Ala	Lys	Val	Gln	Trp	Lys	Val	Asp	Asn	Ala	Leu	Gln	Ser	Gly	
			145					150					155			
	Asn	Ser	Gln	Glu	Ser	Val	Thr	Glu	Gln	Asp	Ser	Lys	Asp	Ser	Thr	
30			160					165					170			
	Tyr	Ser	Leu	Ser	Ser	Thr	Leu	Thr	Leu	Ser	Lys	Ala	Asp	Tyr	Glu	
			175					180					185			
35	Lys	His	Lys	Val	Tyr	Ala	Cys	Glu	Val	Thr	His	Gln	Gly	Leu	Ser	
			190					195					200			
	Ser	Pro	Val	Thr	Lys	Ser	Phe	Asn	Arg	Gly	Glu	Cys				
			205					210				214				

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 300 amino acids
- (B) TYPE: Amino Acid
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

50	Met	Lys	Lys	Asn	Ile	Ala	Phe	Leu	Leu	Ala	Ser	Met	Phe	Val	Phe	
	-23			-20				-15					-10			
	Ser	Ile	Ala	Thr	Asn	Ala	Tyr	Ala	Glu	Val	Gln	Leu	Val	Glu	Ser	
			-5					1					5			
55	Gly	Gly	Gly	Leu	Val	Gln	Pro	Gly	Gly	Ser	Leu	Arg	Leu	Ser	Cys	
			10					15					20			
	Ala	Thr	Ser	Gly	Tyr	Thr	Phe	Thr	Glu	Tyr	Thr	Met	His	Trp	Met	
60			25					30					35			
	Arg	Gln	Ala	Pro	Gly	Lys	Gly	Leu	Glu	Trp	Val	Ala	Gly	Ile	Asn	
			40					45					50			

	Pro	Lys	Asn	Gly	Gly	Thr	Ser	His	Asn	Gln	Arg	Phe	Met	Asp	Arg	
			55					60					65			
5	Phe	Thr	Ile	Ser	Val	Asp	Lys	Ser	Thr	Ser	Thr	Ala	Tyr	Met	Gln	
			70					75					80			
	Met	Asn	Ser	Leu	Arg	Ala	Glu	Asp	Thr	Ala	Val	Tyr	Tyr	Cys	Ala	
			85					90					95			
10	Arg	Trp	Arg	Gly	Leu	Asn	Tyr	Gly	Phe	Asp	Val	Arg	Tyr	Phe	Asp	
			100					105					110			
	Val	Trp	Gly	Gln	Gly	Thr	Leu	Val	Thr	Val	Ser	Ser	Ala	Ser	Thr	
15			115					120					125			
	Lys	Gly	Pro	Ser	Val	Phe	Pro	Leu	Ala	Pro	Ser	Ser	Lys	Ser	Thr	
			130					135					140			
20	Ser	Gly	Gly	Thr	Ala	Ala	Leu	Gly	Cys	Leu	Val	Lys	Asp	Tyr	Phe	
			145					150					155			
	Pro	Glu	Pro	Val	Thr	Val	Ser	Trp	Asn	Ser	Gly	Ala	Leu	Thr	Ser	
			160					165					170			
25	Gly	Val	His	Thr	Phe	Pro	Ala	Val	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	
			175					180					185			
	Ser	Leu	Ser	Ser	Val	Val	Thr	Val	Pro	Ser	Ser	Ser	Leu	Gly	Thr	
30			190					195					200			
	Gln	Thr	Tyr	Ile	Cys	Asn	Val	Asn	His	Lys	Pro	Ser	Asn	Thr	Lys	
			205					210					215			
35	Val	Asp	Lys	Lys	Val	Glu	Pro	Lys	Ser	Cys	Asp	Lys	Thr	His	Thr	
			220					225					230			
	Cys	Pro	Pro	Cys	Pro	Ala	Pro	Glu	Leu	Leu	Gly	Gly	Arg	Met	Lys	
			235					240					245			
40	Gln	Leu	Glu	Asp	Lys	Val	Glu	Glu	Leu	Leu	Ser	Lys	Asn	Tyr	His	
			250					255					260			
	Leu	Glu	Asn	Glu	Val	Ala	Arg	Leu	Lys	Lys	Leu	Val	Gly	Glu	Arg	
45			265					270					275		277	